

# Reproducible Scaling Laws for Contrastive Language-Image Learning

Year: 2022 | Citations: 1107 | Authors: Mehdi Cherti, R. Beaumont, Ross Wightman, Mitchell Wortsman, Gabriel Ilharco

---

## Abstract

Scaling up neural networks has led to remarkable performance across a wide range of tasks. Moreover, performance often follows reliable scaling laws as a function of training set size, model size, and compute, which offers valuable guidance as large-scale experiments are becoming increasingly expensive. However, previous work on scaling laws has primarily used private data & models or focused on uni-modal language or vision learning. To address these limitations, we investigate scaling laws for contrastive language-image pre-training (CLIP) with the public LAION dataset and the open-source OpenCLIP repository. Our large-scale experiments involve models trained on up to two billion image-text pairs and identify power law scaling for multiple downstream tasks including zero-shot classification, retrieval, linear probing, and end-to-end fine-tuning. We find that the training distribution plays a key role in scaling laws as the OpenAI and OpenCLIP models exhibit different scaling behavior despite identical model architectures and similar training recipes. We open-source our evaluation workflow and all models, including the largest public CLIP models, to ensure reproducibility and make scaling laws research more accessible. Source code and instructions to reproduce this study is available at <https://github.com/LAION-AI/scaling-laws-openclip>.